DOCUMENT RESUME

ED 478 982 TM 035 127

AUTHOR	Myers, David; Dynarski, Mark
TITLE	Random Assignment in Program Evaluation and Intervention Research: Questions and Answers.
INSTITUTION	Institute of Education Sciences (ED), Washington, DC.
REPORT NO	NCEE-2003-5001
PUB DATE	2003-06-00
NOTE	10p.
CONTRACT	ED-01-CO-0039-0002
PUB TYPE	Reports - Descriptive (141)
EDRS PRICE	EDRS Price MF01/PC01 Plus Postage.
DESCRIPTORS	Control Groups; *Educational Research; *Evaluation Methods; *Experimental Groups; *Intervention; *Program Evaluation
IDENTIFIERS	*Randomized Experiments

ABSTRACT

This booklet contains questions and answers about random assignment in program evaluation and intervention research. The main purpose of program evaluation research in education is to determine whether programs help the students they are designed to serve and whether new ideas for education programs still under development are worthy of extension to a wider selection of schools and settings. Researchers use random assignment in impact studies to form two statistically equivalent groups of participants in the most objective way possible. Random assignment procedures vary according to the program being tested. The main advantage of using random assignment is a "one-difference-only" feature to help determine whether changes are the result of the intervention or program. Other approaches can be used to study the effects of a new program, but random assignment is the best approach for discerning whether or not a new school policy or program has an effect over and above the effect of the standard policy or program and how big that effect might be. (SLD)

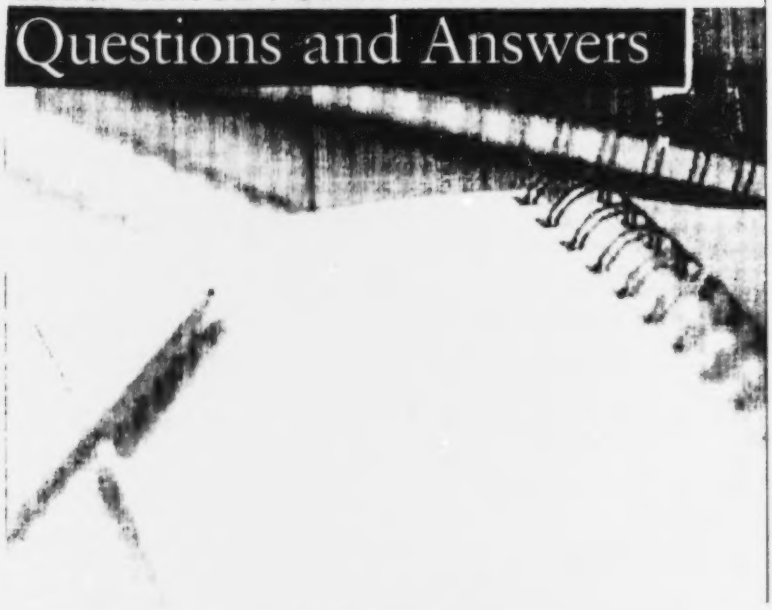
Reproductions supplied by EDRS are the best that can be made
from the original document.

TM TMTR 0178

ED 478 982

Random Assignment

Random Assignment in Program Evaluation and Intervention Research: Questions and Answers



U.S. Department of Education
Institute of Education Sciences
National Center for Education Evaluation
and Regional Assistance

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

TM035127

STAR Data Entry Database: DIP Record: 25102

Title	Random Assignment in Program Evaluation and Intervention Research: Questions and Answers.
Orga	Institute of Education Sciences
Orgb	National Center for Education Evaluation & Regional Assistance
Orgc	AAF
Author	Myers, D. and Dynarski, M.
Sponsor	IES
Idnum	
Pub Date	2003-06-00
Copies	1
Notes	
Date Ordered	
CH Assigned to	TM
Date Sent	2003-07-01
Dip Num	25102
Stamp	JMILLE |D 2003 06 30 |T
Update	JMILLE |D 2003 06 30 |T

-- INSERT MODE

Command: Recall

Help DAtabase Create Recall Next PRevious COpy IGnore Save Update DElete ERase
Options Display Print Write Index ENTER SEARCH GLOBAL MENU

U.S. Department of Education
Rod Paige
Secretary

Institute of Education Sciences
Grover J. Whitehurst
Director

June 2003

By: David Myers and Mark Dynarski
Mathematica Policy Research

This project was supported under contract number ED-01-C0-0039 (0002), as administered by the Institute of Education Sciences, U.S. Department of Education. This publication is in the public domain. Authorization to reproduce it in whole or in part for educational purposes is granted.
The full text of this publication is available at http://www.ed.gov/offices/IES/NCEE/qa.html.

BEST COPY AVAILABLE

Random Assignment in Program Evaluation and Intervention Research: Questions and Answers

What is the purpose of education program evaluation?

The basic purpose of program evaluation research is to test (1) whether education programs help the students they are designed to serve and (2) whether new ideas for education programs still under development are worthy of extension to a wider selection of schools and settings. The term "program" refers to a specific set of education practices or interventions that are thought to have an impact on a given set of education outcomes.

For example, imagine that a new math program appears to show promise in improving students' math skills. But before teachers, school administrators, and education policymakers make that program a permanent part of the curriculum, they will want to know whether the program actually helps students learn mathematics better than the usual program. To answer this question, a local school district, a state Department of Education, or the U.S. Department of Education may commission a study to evaluate the new program. This effort would involve

setting up the program in certain schools or classrooms to examine how students in the program are doing in math compared to other students not in the program. In technical terms, this examination is called an "impact study" because its purpose is to determine the impact, or the effect, of the program on students' math skills. The impact is defined as the difference between the math skills of students in the new program and the math skills of students not in the program. The most reliable and accurate way to conduct an impact study is through a technique known as random assignment.

What is random assignment?

Researchers use random assignment in impact studies to form two statistically equivalent groups of participants in the most objective way possible. The term "participants" may refer to students, teachers, classrooms, or schools.

Random assignment procedures vary according to the program being tested. In most situations, the basic process of randomly assigning participants to an education intervention and a control condition are similar. First, a list of participants is created. The list of participants is then randomly assigned to conditions. Some of the participants are assigned to the program condition, and the other participants are assigned to the control condition. The actual number of participants assigned to the program and control conditions, carefully thought out by the evaluator, is based on statistical considerations such as the "size" of the impact one wants to detect and the chances of observing large impacts by chance.

For instance, in a random assignment impact study of our hypothetical math program, classrooms would be the participants, and they would be selected in a purely random way to be part of one of two conditions. One group of classrooms would have the new math program in its curriculum—this is known as the "intervention" or "program" condition. The other group of classrooms would continue in the school's regular math program—this is known as the "control" condition.

On the other hand, if a school system wants to try out a new method of professional development for teachers to see whether it changes teaching practices in a way that helps students score higher on achievement tests, teachers would be the participants who are randomly assigned to the intervention or the control condition. The former would be exposed to the new professional development program, and the latter would be offered the regular program.

In yet another example of random assignment, suppose a state wants to assess different comprehensive school reform strategies as the avenue to better test scores. In this case, schools would be the participants randomly assigned to the various reform conditions (i.e., interventions) or to the control condition. In other words, the reforms would be put into place in some schools, but not in others, and random assignment would be used to select the schools to implement the reforms.

The assignment process works much the same as a lottery, so each participant—whether a student, a classroom, a teacher, or a school—has the same chance of ending up in the intervention or the control group. For instance, in our hypothetical math program, imagine that each classroom is given a number on a slip of paper. The

slips of paper are then placed in a bin and shuffled, and half the numbers are pulled from the bin. Classrooms with these numbers are assigned to the new math program (intervention condition), and classrooms with numbers remaining in the bin continue in the regular program (control condition). In practice, the assignment process is performed by a computer.

Because classrooms are assigned to one condition or the other wholly at random, the chances of a more or less capable, or a more or less motivated classroom of students ending up in one condition or the other are the same. So the only difference between the two conditions is the type of math instruction they are getting. This "one-difference-only" feature is the main advantage of using random assignment in an impact study, because it means that if the math skills of the students in the new program improve more than the math skills of the students in the regular program, we can be almost completely certain that the improvement occurs because of the program, not because of student ability or motivation or other factors that might influence their achievement.

Why random assignment instead of another approach?

Researchers sometimes use approaches other than random assignment in an impact study to determine the effect of a new program. One example of an alternative approach is a simple "pre-post" comparison. Here, using our new math program again, the study would compare students' math test scores *before* they participated in the new program with their scores *after* they participated. Any difference in the scores is assumed to be a result of the program. Simple pre-post comparisons do not support the conclusion that the changes over time in student performance are due to the new program, because students and their social and educational environment change naturally over time. So it is impossible to isolate changes due to a program from these natural changes that may have occurred while the program was being tried out.

Another example is known as the "matched comparison" approach. Here, researchers would use statistical techniques to find students who did not participate in the new program, but who are similar to (i.e., who "match") those who did. The math scores of participating students would be compared with the math scores of nonparticipating students, and again, any difference in scores is assumed to be due to the program.

The matched comparison approach can have low validity depending on how much information researchers are able to use to match groups. For example, even though it can help researchers start out with two groups of similar classrooms that, as in random assignment, differ in that one has implemented the new program and the other has not, the two groups of classrooms are similar *only with regard to* the characteristics used by the researchers to make the match. For example, the classrooms may be similar in size, race/ethnicity, and family income, but they may differ in terms of teacher motivation, quality of teaching, expectations for students, and student motivation. So if we do see an improvement in student performance in the classrooms that implemented the new program, it is possible that this improvement is the result of pre-existing factors *other than* the interventions, like

teacher motivation, or the result of these factors *combined with* the interventions.

Even when using carefully constructed comparison groups, a comparison of impacts measured using random assignment and impacts of the same intervention measured using carefully constructed comparison groups often shows that the results are different, sometimes much different. For example, a recent random assignment study of dropout prevention programs showed that some types of interventions were effective, but when a matched comparison group design was used instead of a random assignment design, the interventions appeared to be highly ineffective.

Comparison groups that are not carefully constructed—such as groups consisting of participants who stop participating, participants who participate infrequently, students who express interest but elect not to participate, or students who are eligible and do not express interest—have the lowest validity, as their decisions indicate that they differ from participants who continued with the program. Comparison groups constructed using these "samples of convenience" are likely to fall well short of the desired standard of rigor. Because of these limitations, random assignment plays a leading role in program evaluation, because it is the only approach that can make us confident that what we see after a new program is put in place is due to that program alone, and not to the many other factors that are in play in schools and in the lives of students. In other words, random assignment is the best approach for discerning whether a new school policy or program has an effect over and above the effect of the standard policy or program and how big that effect may be.

Is participation in the evaluation voluntary?

In some cases, participation in a study may be voluntary, and in others, it may be required as a consequence of a school district or school trying out a new educational practice. For example, if researchers are examining the effectiveness of an afterschool program, students' participation in the study may be voluntary—parents can decide whether their child will attend the program. But when researchers are examining the effectiveness of different teaching methods, such as variations on how technology may be integrated into the classroom, schools may randomly assign one variation or another to entire classrooms. In this case, participation is "required" in that students may not have an option to change classes. Similarly, school districts that want to test alternative whole-school reforms may decide that all schools should participate in the evaluation. In all cases, however, the evaluator should work closely with school districts, schools, teachers, students, and parents to ensure that everyone has a sound understanding of the importance of the study and the study procedures.

Does the intervention group need to have the same number of participants as the control group?

Some variation in the number of participants in the two groups can be accommodated. For example, a study of the impact of offering a new type of professional development for teachers might have 200 teachers participate in the new program and 100 teachers participate in their regular professional development activities.

Is random assignment a fair way to select participants?

Random assignment is fair, because it gives all participants an equal chance of being selected for the program. Personal factors play no role in whether a participant is selected. Programs often have more applicants than they are able to serve, and using random assignment is the same as "picking names from the hat" or using a lottery to allocate limited program spaces.

Do you have to change the process for recruiting participants?

In some contexts, recruiting efforts may need to be intensified because random assignment requires more participants (students, classrooms, or schools) than a new program is intended to serve. For example, if a school district wanted to institute an education reform program, and it had adequate funding to do so in five schools, and if it wanted all schools participating in the study of the reform program to be volunteers, it would need to recruit more than five schools at the outset so that there would be enough schools to form an intervention group and a control group. Similarly, if a school district wanted to determine the effect of an afterschool program, the program staff would need to recruit about twice as many students as could be accommodated in the program. Half the students would participate in the afterschool program, and the other half would be in the control condition. However, when a school district is free to assign classrooms or schools to a new program without asking for volunteers, recruitment is not necessary. Likewise, when a new program naturally generates far more volunteers than can be served, no special recruitment process is necessary.

What happens to participants in the control condition? Can they receive other kinds of services?

Participants assigned to the control condition can have access to any services not provided by the program or intervention being studied. For instance, parents of a student assigned to a control condition in a study of an afterschool program could seek out some other type of afterschool care for their child. It is important for evaluators to collect data about the services that control group members receive so that they can better interpret the evaluation's findings.

Will participants in the control group ever have an opportunity to receive the new services?

Assuming the new program proves effective, which we do not know prior to the evaluation, participants in the control condition can become eligible for the new program as soon as the study is complete. Having participants in the control condition be first on the list for enrollment in the new program once the study is complete can be an incentive for participation.

Are there situations in which random assignment is or isn't appropriate?

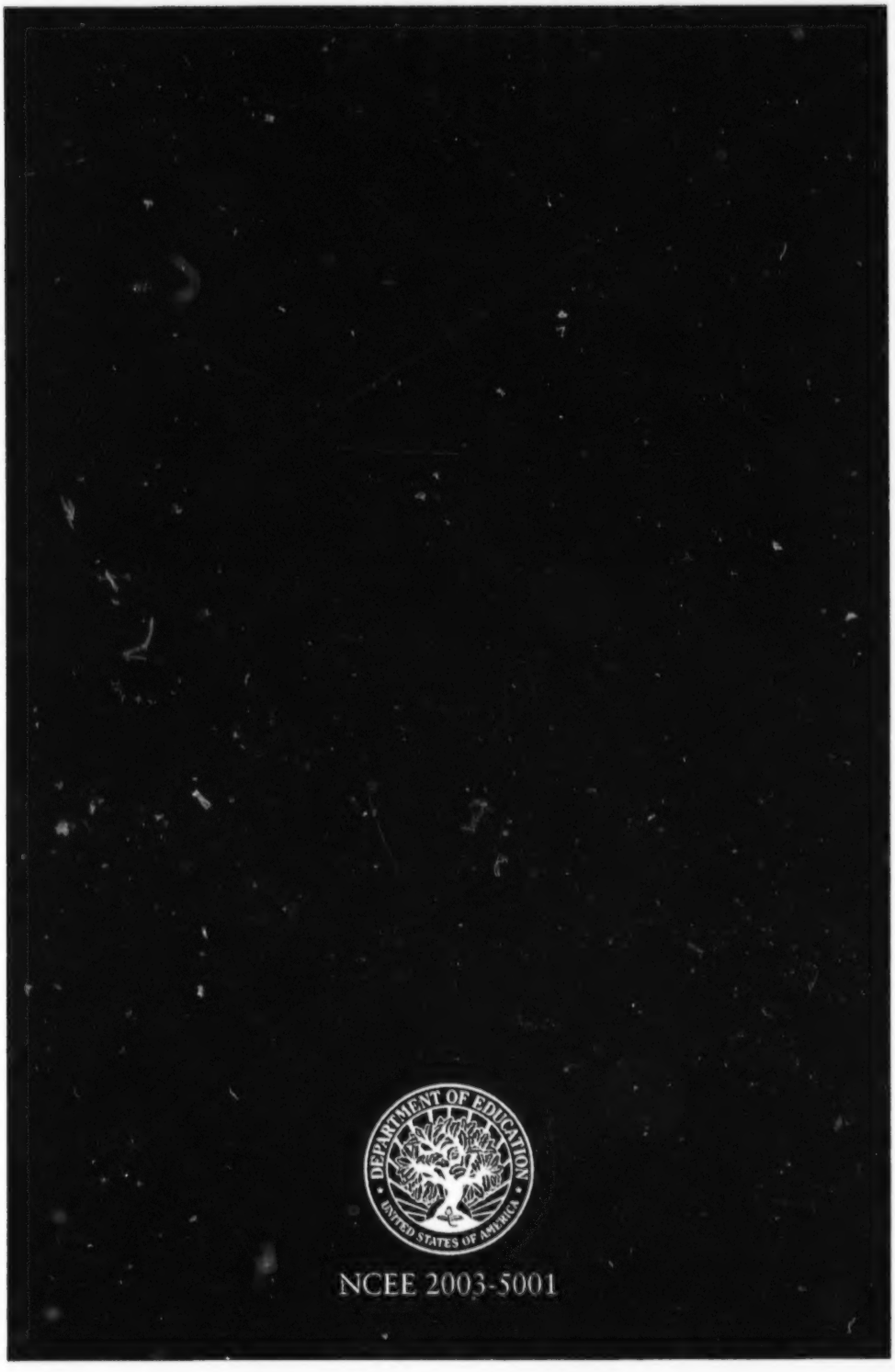
If the consequences of random assignment create a situation that is potentially harmful to some students who might end up in the control condition, it would be appropriate to exempt those students from the research. For instance, in planning for an impact study of the effects of Head Start, the federal early childhood program for children in poverty, evaluators might decide to exempt children with the highest need from participation in random assignment. However, both the exemption procedures and the circumstances under which students are exempted should be firmly established before the study begins. In turn, study results are not applicable to students who are exempted.

What happens if a participant "drops out" of the program after being randomly assigned to the intervention group?

Participants who drop out of a program or intervention will be viewed as if they were still part of the study. That is, researchers will continue to collect followup data from all participants in the study regardless of whether they drop out. This allows researchers to assess the extent to which dropping out influences the effects of the program. It also prevents the problem of a new program appearing to be more effective than a control condition because participants making less progress drop out of the new program at higher rates than they drop out of the control condition.

What would you gain by participating in a program evaluation?

Participating in an education program evaluation has many advantages. First, it gives schools, teachers, and students an opportunity to help policymakers learn what education practices or interventions are particularly effective. Right now, we have very little information on what works and what doesn't. Only with the assistance of schools, teachers, and students can those who design education programs really learn what will make education more equitable and efficient for all children. Second, study results can help school districts, for example, make better decisions about the education practices they regularly use and about new practices that they may use on a larger scale in the future. Third, participating in an evaluation may give school districts and schools access to resources that might not otherwise be available. Finally, participating school districts, schools, and students will also be the first to reap the benefits of new and possibly innovative programs if the findings indicate that the programs are effective. The 50 percent chance of experiencing a new, exciting, and potentially important educational innovation that comes though participating in a randomized study is often much better odds than are available for those students and schools that don't participate at all and thus don't have access to the innovation.



NCEE 2003-5001

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

[ ] This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

[X] This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").